DOCUMENT RESUME

ED 276 259 FL 016 211

AUTHOR Birdsong, David
TITLE Empirical Impediments to Theories of Second Language Acquisition.
PUB DATE Apr 86
NOTE 11p.; Paper presented at the Kentucky Foreign Language Conference (Lexington, KY, April 24-26, 1986).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adult Learning; *Data Interpretation; *Grammar; *Language Universals; Learning Processes; *Linguistic Theory; *Research Methodology; *Second Language Learning

ABSTRACT

A recent wave of linguistic research has attempted to demonstrate empirically that in learning the syntax of a second language, adults re-access their universal grammar (UG). However, the conspicuous lack of success of second language learners has caused researchers to seek evidence in experimental data. The most expedient source of this data is grammaticality judgments, and they present problems of reliability and validity. More trustworthy elicitation techniques have been developed, but grammaticality judgments are still being used with few methodological safeguards. The most serious barrier to the development of a "principled theory" of second language learning appears to be reliance on unprincipled data. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED276259

EMPIRICAL IMPEDIMENTS TO THEORIES OF SECOND LANGUAGE ACQUISITION

David Birdsong, Department of Romance Languages and Literatures,
University of Florida, Gainesville, FLORIDA 32611

Paper presented at the 1986 Kentucky Foreign Language Conference,
University of Kentucky, Lexington, KENTUCKY April 24-26, 1986

INTRODUCTION

In current second language research, there is much ado about the quest for a "principled theory" to account for adults' acquisition of non-native syntax (see, e.g., Flynn 1985b). It is understood that such a theory should emerge from a formal model of language acquisition, and should be capable of making testable predictions of, and providing adequate explanations for, the linguistic behavior of learners. For first language (L1) acquisition, a stellar example of such a theory is Steve Pinker's (1984) Language Learnability and Language Development. For second language (L2) acquisition, no such fleshed-out version yet exists. [1] However, there is agreement among a considerable number of researchers as to the point of departure for a legitimate L2 acquisition model. This point of departure turns out to be the same as that posited by Chomsky (1980, inter alia) for L1 acquisition, namely, Universal Grammar.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

David Birdsong

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

1 See, however, Eubank (1986) for extensive discussion of UG's applicability to L2 and L1 acquisition of German.

Universal Grammar (UG) has been proposed as a solution to what is often referred to as the logical problem of first language acquisition: how is the child able to produce more sentences than have been available in input, and how is (s)he able to puzzle out the complexities of the mother tongue in so brief a period of exposure? The answer resides in the postulate that the child is genetically endowed with a set of cognitive principles which are specific to the domain of language learning, and which are identified as Universal Grammar:

> Universal Grammar is taken to be the set of properties, conditions, or whatever that constitutes the 'initial' state of the language learner, hence the basis on which knowledge of language develops. (Chomsky, 1980: 69).

UG defines a limited range of possible grammars of natural languages, and thus reduces the learner's pure guesswork in comprehension and production of novel utterances. The operation of UG in language acquisition is said to involve learners' setting of parameters and their distinguishing between core and peripheral grammar (see below).

A recent wave of researchers has attempted to demonstrate empirically that, in learning the syntax of a second language, adults re-access UG. White (1985b) examines L2 acquisition of the UG parameters of SUBJACENCY (which deals with the number and type of bounding nodes across which sentence elements can be extracted) and PRO-DROP (which allows for omission of the subject pronoun, free inversion of subject and verb in declaratives, and "that-trace" effects, where a subject is extracted out of a clause containing a complementizer). Her claim is that L2 acquisition involves parameter-setting (the learner must determine, for example, whether the target language is a PRO-DROP language), except that with a second language the learner re-sets parameters that differ from the native tongue. A similar line of research is undertaken by Flynn (1985a), working with the parameter of PRINCIPAL BRANCHING DIRECTION. Mazurkewich (1985) has sought to demonstrate that no matter what the native language of the L2 learner, unmarked (or core grammar) structures such as Direct Object Passive (e.g. 'The ball was thrown by John') will be acquired before marked (or peripheral) structures such as Dative Passives (e.g., 'The ball was thrown to John' or 'John was thrown the ball'). Felix (1984), working with native speakers of German, illustrates seven structural constraints of UG with grammatical and ungrammatical target language (English) sentences; these sentences have no equivalent in German and have not been explicitly taught to the ESL subjects. Among the constraints illustrated are SUPERIORITY EFFECTS (e.g. '*I don't know what who did' vs 'I don't know who did what'), EXCEPTIONAL CASE MARKING (e.g. '*Fred was easy for Bob to

expect to come to the party' versus 'Fred was easy for Bob to persuade to come to the party'), and the SPECIFIED SUBJECT CONDITION (e.g., '*Who did the man see John's pictures of?' versus 'Who did the man see pictures of?'). ESL subjects, when asked if the representative sentences were grammatical or ungrammatical, performed better than chance overall. Their performance on this task is attributed to their accessing of the relevant constraints in Universal Grammar.

Despite the significant contributions of such lines of research, and despite the increasing acceptance of the role of UG in first language acquisition, there are serious obstacles to its application to second language acquisition. The present paper offers a number of perspectives on the interplay--and frequent mismatches--of empirical linguistic data and L2 acquisition theory. Though many of the observations will be applicable to quite a variety of current theories of second language acquisition, emphasis will be on the prevailing UG paradigm. Moreover, since much of the core data within this and other paradigms has been in the form of learners' grammaticality judgments, their limitations and proper interpretation will be discussed. Time constraints will, unfortunately, limit discussion to fairly narrow domains of learner output data, to the complete exclusion of input data.[2]

END-PRODUCT DATA AND THE UG/L2 PARADIGM

A major impediment to second language acquistion theory based on Universal Grammar is the conspicuous lack of uniform success among L2 learners. We will recall that UG theory, as applied to L1, is above all a response to the logical problem of uniform rapid progress and success among children learning their mother tongue. No such miracle characterizes L2 acquisition. The end products of adults' attempts at L2 acquisition range from virtual zero mastery to native-like mastery, with the distribution of

[2] For detailed discussion of input and output data in L2 acquisition, see Bley-Vroman (1986a,b). Any L2 acquisition model that appeals to the essential identity of L2 and L1 acquisition mechanisms (and the avatars are many) must maneuver around a central empirical fact, namely, that the nature of the linguistic input available to the child first language learner is demonstrably different, both quantitatively and qualitatively, from that which is available to the adult second language learner. Even the spersest of catalogulngs of the differences would include adults' access to linguistic input via the written as well as the oral code, the heterogeneity of input types in both classroom and naturalistic environments, adults' acquaintance with semantic subtleties specific to speech situations, etc., not to mention well-documented differences in L2 learning styles.

learners falling roughly within a normal bell curve (see Scovell 1985). It is fair to say that, relative to the uniform success of L1 children, L2 adults display varying degrees of failure. A UG account, therefore, is not motivated on the same empirical grounds for L2 as for L1.[3]

It may be argued, then, that at best a UG model is not necessary to account for generally unsuccessful acquisition of an L2; at worst, it is inappropriate to invoke a paradigm of linguistic behavior that is essentially phylogenetic and deterministic in the face of end-product data that are essentially idiosyncratic and indeterminate. Additional arguments in favor of this position may be found in Bley-Vroman (1986a).

A counterargument to the above might run as follows. Though end-product data suggest that (most) adult L2 learners do not apply UG in the manner suggested by Chomsky for L1 children, this is not to say that adults cannot access UG. If we assume that adults have ability to access UG, then we open the door to quite a variety of logical possibilities, the permutations depending on whether: (a) some or all adults can access UG; (b) UG is accessed fully or partially; (c) UG is applied fully or partially. Note that no claim about the structure or content of UG is being made; rather, an enhanced range of possibilites of access and application is being considered. The appeal of such a weakened version of UG's role in L2 acquistion is its flexibility, which would allow one to account for the variability of success attested among learners of a second language.

The dangers inherent in such an argument are obvious. Any theory, however preposterous or plausible, could account for data if applied on an ad hoc (learner-by-learner, structure-by-structure) basis. Moreover, the weakened version offers a tempting invitation to adduce validating evidence via post hoc ergo propter hoc reasoning.

Nevertheless, one should not overlook a compelling feature of the weakened version, namely the premise that it is possible for L2 learners to access UG. Since application of UG assumes its availability, establishing the validity of this premise becomes a logical necessity. Further, certain L2 acquisition theories (e.g. Felix' [1985b] competition model) are built on the argument that UG is indeed accessible, but that other cognitive structures compete with it when the learner is confronted with L2 learning tasks. It is not surprising, then, that considerable research has been directed toward demonstrating that adults can access UG, leaving open the issue of whether they can or do apply it to L2 acquisition.

---

[3] It is nonetheless true that, as in L1 acquisition, the range of L2 learners' utterances can exceed that found in their input. This fact is not uniquely accounted for, though, by the UG paradigm. For further discussion, see Felix (1985b) and Bley-Vroman (1986a).

GRAMMATICALITY JUDGMENTS IN THE UG/L2 PARADIGM

Much of the evidence for L2 learners' potential for accessing UG has been in the form of grammaticality judgments; that is, experimental elicitations of learner intuitions for sentences compatible with, or in violation of, UG constraints. Grammaticality judgments are more expediently collected than natural production data, while ostensibly reflecting the systematic nature of learner "competence" rather than erratic "performance". Judgments of grammaticality, according to White (1985a: 37), constitute "a very important source of data, a means to tap learner intutions about the L2." The validity of such data and their role in current grammatical theory are discussed at length in Newmeyer (1983, ch. 2), while a thorough review of L1 and L2 research based on linguistic judgments appears in Chaudron (1983). Recent research within the UG/L2 paradigm has revealed that learners' intuitions for the grammaticality or ungrammaticality of unfamiliar sentence types in a non-primary language are accurate, on the average, at a better-than-chance rate. Such performance suggests a basis for judgments other than knowledge of their native or target language. It is argued that the basis for judgment is Universal Grammar, with the conclusion that UG can in fact be accessed by L2 learners.

While better-than-chance accuracy on judgment tasks is provocative, it is unwise to view such performance as decisive. Our cautiousness derives from consideration of judgment variability and, at the root of this variability, certain attested psycho/cognitive factors among subjects in judgment tasks.

Salient examples of judgment variability are found in the Felix (1985a) study. Forty-eight native speakers of German were asked to judge 24 English sentences; seven unfamiliar structural types were represented by four exemplars each. One-half of the stimuli sentences were grammatical, half were ungrammatical. Overall, approximately 69% of stimuli sentences were correctly judged; 80.5% of the ungrammatical sentences were properly rejected, while 57.5% of grammatical sentences were properly accepted. Though these figures seem quite impressive, the variability behind the numbers hinders straightforward interpretation. For example, across structural types, correct responses range from 58.6% for CONTROL/EXCEPTIONAL CASE MARKING sentences (see above) to 91% for CASE FILTER EFFECT sentences ('John seems to love Mary' versus '*Mary seems John to love'); within sentence types, correct responses range from 29.2% proper acceptance to 95.8% proper rejection of PARASITIC GAP sentences ('a person that they spoke to because they admired' versus '*a

person that they spoke to because admired them'). After extensive discussion, Felix' response to such variable effects is the claim that learners access various UG principles at different times along a developmental sequence. However, such an argument does not square with data from adult English native speakers, who presumably have completed their developmental sequence: within-type correct response rates for these "control" subjects ranged from 0% to 100% for PARASITIC GAP exemplars. Such results may be methodological artifacts (see discussion below).

As for the difference in accuracy between stimulus sentences which are grammatical (57.5% correctly judged) and those which are ungrammatical (80.5% correctly judged), Felix maintains:

> Considering the theoretical status of UG within the logical problem of language acquisition, this is, in fact, what we would expect. The primary function of Universal Grammar is to provide the child with information as to which structures are *un*grammatical, since there is no other negative evidence available to the language learner. If a given construction violates [a] principle[s] of UG, the child will know for sure that this construction must be ungrammatical. If, in contrast, a structure does *not* violate UG, it may still be unacceptable for other reasons. In other words, UG is an unambiguous source of information for *un*grammaticalities, but only a subsidiary source for identifying grammatical sentences. (1985a: 13)

This appealing argument, for whatever its ultimate truth, fails to respond to a curious asymmetry in the response data: non-native subjects supplied the judgment, "ungrammatical" about 61.5% of the time, while judgments, "grammatical," constituted only 38.5% of the responses. Subjects faced with the task of making grammaticality judgments may tend naturally to supply the judgment, "ungrammmatical," more often than the judgment, "grammatical" (see Chaudron 1983: 364); thus the accuracy rate for stimuli sentences which are indeed ungrammatical is predictably greater than that for grammatical sentences. Returning to the Felix data, when one compares the percentage of judgments "grammatical" (38.5%) to correct responses on grammatical stimuli (57.5%), the observed freqency of correct responses is superior to the expected frequency of correct responses by 19%; a similar comparison with judgments "ungrammatical" (61.5%) and observed accuracy (80.5%) yields an identical 19% better-than-expected rate of accuracy. Thus it is inappropriate to claim that subjects' performance on ungrammatical items was "superior" to their performance on grammatical items. Such a claim ignores a fact of experimental performance, namely, subjects' bias toward supplying judgments "ungrammatical."

One should not overlook, however, certain variability data which might support Felix' position concerning judgments of ungrammatical sentences. The 57.5% accuracy rate for grammatical stimuli embodies a staggering standard deviation across items of 17.6, while the 80.5% figure for ungrammatical sentences reflects a relatively small standard deviation of 10.5. Thus, compared to grammatical sentences, one may look with considerably more confidence on the accuracy rate for ungrammatical sentences. All things considered, though, the types and magnitude of variability in the Felix study make unambiguous interpretation next to impossible.

Numerous sorts of variability in metalinguistic judgment tasks are documented in Chaudron (1983), as are methodological sources of "noisy" judgment data. Subtle psychological and cognitive factors may also generate variabilty in judgments. For example, Armstrong, Gleitman, and Gleitman (1983) found that even for well-defined, nominal (binary) categories such as "even number," "odd number," "plane geometry figure," "female," "male," etc., subjects consider certain exemplars of a given category more representative or "better" members than others. In this study, subjects viewed numbers like 806 as "less even" than 4; trapezoids and ellipses are considered not as plane geometrical as squares; a princess is deemed less representative of "female" than mothers and ballerinas. This suprising effect was obtained not only in judgment tasks, but over a variety of experimental conditions (e.g., in timed verification tasks, subjects took longer to recognize the truth or falseness of statements like, "703 is an odd number," than statements like, "9 is an odd number"). Arguably, the variability attested for judgments of everyday nominal categories might apply to the category "grammatical/ungrammatical sentence." We should not be surprised, then, if some exemplars of a given grammatical constraint are judged more (or less) grammatical than others. One among many illustrations of such inter-item variability is available in judgments of violations of the (universal) Coordinate Structure Constraint. Informal surveys of learners and native speakers of English have revealed divergent rejection rates across tokens such as, "*Who did you see Mary and?", "*What is it raining cats and?", and "*What did the grandfather clock stand between the bed and?" (see Ross, 1979). Note that the inter-item variability just described is a fact of experimental performance, and may thus be considered independently of rating variability attributable to "fuzzy" grammatical constraints (see Mohan 1977, Birdsong 1984). Analysis and interpretation of judgment data therefore critically hinge on the fundamental comparability of items to be judged; at issue is nothing less than the internal validity of the elicitation instrument.

In a 1981 study entitled, "The non-uniqueness of linguistic intuitions," Carroll, Bever, and Pollack address psychological sources of intra-subject variability. They find that it is possible to induce changes in native speakers' language judgment patterns by manipulating physical elements of the experimental setting. Patterns of response for individual subjects were found to differ merely as a function of the presence or absence of a mirror in the room where the experiments took place! Once again, legitimate use of judgment data becomes a question of procedural control and caution.

Yet, even when elaborate procedural safeguards are taken, subjects may perform in unexpected ways. In a study by Birdsong et al. (1985), learners of French as a foreign language were instructed to make grammaticality judgments solely on the basis of syntactic criteria, specifically, 'Words Out of Place, Missing Words, or Too Many Words'. A post-test of validity was carried out with deviant sentences which varied minimally by the morphology of their complementizers. Paired sentences such as, '*Voilà les enfants QUI Pierre a joué avec eux hier' versus '*Voilà les enfants QUE Pierre a joué avec eux hier'; '*Voilà les enfants QUE Pierre a joué avec hier' versus '*Voilà les enfants LESQUELS Pierre a joué avec hier', were differentially judged, despite explicit instructions to focus on syntactic features.

CONCLUSION

Faced with a lack of uniform success among L2 acquirers, researchers seek evidence of the role of UG from experimental data. The most expedient source of such data, grammaticality judgments, proves to be rife with problems of reliability and validity. While data derived from novel, sophisticated elicitation techniques appear more trustworthy (see, for example, Finer and Broselow 1986), the use of grammaticality judgments, with few if any methodological safeguards, continues. It is ironic that perhaps the most palpable impediment to the development of a principled theory of L2 acquisition is reliance on unprincipled data.

REFERENCES

Armstrong, S.L., L. Gleitman, H. Gleitman (1983) What some concepts may not be. Cognition, 13, 263-308.

Birdsong, D. (1984) Prenominal past participles in French. In P. Baldi, ed. Papers from the XIIth linguistic symposium on romance languages (Current Issues in Linguistic Theory, vol. 36), 37-50. Amsterdam: John Benjamins.

Birdsong, D., C. Johnson, J. McMinn, P. Ingmundson (1985) Universals versus transfer revisited. ERIC ED 255 061.

Bley-Vroman, R. (1986a) The logical problem of foreign language learning. To appear in Linguistic Inquiry.

Bley-Vroman, R. (1986b) Hypothesis testing in second language acquisition theory. To appear in Language Learning.

Carroll, J., T. Bever and C. Pollack (1981) The non-uniqueness of linguistic intuitions. Language, 57, 368-382.

Chaudron, C. (1983) Research on metalinguistic judgments: a review of theory, methods, and results. Language Learning, 33, 343-377.

Chomsky, N. (1980) Rules and representations. New York: Columbia University Press.

Eubank, L. (1986) Formal models of language learning and the acquisition of German word order and negation by primary and non-primary language learners. Unpublished doctoral dissertation, University of Texas at Austin.

Felix, S. (1984) UG-generated knowledge in adult second language acquistion. Unpublished manuscript.

Felix, S. (1985) Transfer of mental structures: more evidence on competing cognitive structures. Unpublished manuscript.

Finer, D.L. and E.I. Broselow (1986) Second language acquisition of reflexive binding. Paper presented at the sixteenth New England Linguistic Society meeting.

Flynn, S. (1985a) A universal in L2 acquisition based on a PBD typology. In F. Eckman and L. Bell, eds., Universals in second language acquisition, pp. 75-87. Rowley, Mass.: Newbury House.

Flynn, S. (1985b) Principled theories of L2 acquisition. Studies in Second Language Acquisition, 7, 99-107.

Mazurkewich, I. (1985) Syntactic markedness and language acquisition. Studies in Second Language Acquisition, 7, 15-36.

Mohan, B.A. (1977) Acceptability testing and fuzzy grammar. In S. Greenbaum, ed., Acceptability in language. (Contributions to the Sociology of Language, vol. 17), pp. 133-148. The Hague: Mouton.

Newmeyer, F. (1983) Grammatical theory: its limits and its possibilities. Chicago: University of Chicago Press.

Pinker, S. (1984) Language learnability and language development. Cambridge, Mass.: Harvard University Press.

Ross, J.R. (1979) Where's English? In C.J. Fillmore et al., eds., Individual differences in language ability and language behavior, pp. 127-163. New York: Academic Press.

Scovell, T. (1985) The insufficiency of comprehensible input and low affect: biological constraints on non-primary language acquisition. Paper presented at the annual meeting of Texas Teachers of English to Speakers of Other Languages, November 8, 1985.

White, L. (1985a) Is there a 'logical problem' of second language acquisition? TESL Canada Journal, 2, 29-41.

White, L. (1985b) The 'pro-drop' parameter in adult second language acquisition. Language Learning, 35, 47-62.